# ENTERED

RECEIVED

AUG 1 1 2003

TECH CENTER 1600/2900

OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/036,729

DATE: 08/05/2003
TIME: 10:47:04

Input Set : **N:\EBONY'S\US10036729.raw.txt**
Output Set: **N:\CRF4\08052003\J036729.raw**

**SEQUENCE LISTING**

```
 1 (1) GENERAL INFORMATION:
 2      (i) APPLICANT: Middeldorp, Jaap Michiel.
 3     (ii) TITLE OF INVENTION: Peptides and nucleic acid sequences
 4                             related to the Epstein-Barr virus.
 5    (iii) NUMBER OF SEQUENCES: 22
 6     (iv) CORRESPONDENCE ADDRESS:
 7          (A) ADDRESSEE: Akzo-Nobel Patent Department
 8          (B) STREET: 1300 Piccard Drive, Suite 206
 9          (C) CITY: Rockville
10          (D) STATE: Maryland
11          (E) COUNTRY: USA
12          (F) ZIP: 20850
13      (v) COMPUTER READABLE FORM:
14          (A) MEDIUM TYPE: Floppy disk
15          (B) COMPUTER: IBM PC compatible
16          (C) OPERATING SYSTEM: PC-DOS/MS-DOS
17          (D) SOFTWARE: Patentin Release #1.0, Version #1.25
18     (vi) CURRENT APPLICATION DATA:
19          (A) APPLICATION NUMBER: US/10/036,729
20          (B) FILING DATE: 21-Dec-2001
21    (vii) PRIOR APPLICATION DATA:
22          (A) APPLICATION NUMBER: 08/415,838
23          (B) FILING DATE:
24   (viii) ATTORNEY/AGENT INFORMATION:
25          (A) NAME: Gormley, Mary E.
26          (B) REGISTRATION NUMBER: 34,409
27 (2) INFORMATION FOR SEQ ID NO: 1:
28      (i) SEQUENCE CHARACTERISTICS:
29          (A) LENGTH: 538 base pairs
30          (B) TYPE: nucleic acid
31          (C) STRANDEDNESS: double
32          (D) TOPOLOGY: unknown
33     (ii) MOLECULE TYPE: DNA (genomic)
34     (vi) ORIGINAL SOURCE:
35          (A) ORGANISM: Epstein-Barr virus
36     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
37 CATGATGGCA CGCCGGCTGC CCAAGCCCAC CCTCCAGGGG AGGCTGGAGG CGGATTTTCC      60
38 AGACAGTCCC CTGCTTCCTA AATTTCAAGA GCTGAACCAG AATAATCTCC CCAATGATGT     120
39 TTTTCGGGAG. GCTCAAAGAA GTTACCTGGT ATTTCTGACA TCCCAGTTCT GCTACGAAGA     180
40 GTACGTGCAG AGGACTTTTG GGGTGCCTCG GCGCCAACGC GCCATAGACA AGAGGCAGAG     240
41 AGCCAGTGTG GCTGGGGCTG GTGCTCATGC ACACCTTGGC GGGTCATCCG CCACCCCCGT     300
42 CCAGCAGGCT CAGGCCGCCG CATCCGCTGG GACCGGGGCC TTGGCATCAT CAGCGCCGTC     360
```

C--> 19
C--> 20

RAW SEQUENCE LISTING                    DATE: 08/05/2003
PATENT APPLICATION: **US/10/036,729**    TIME: 10:47:04

Input Set : **N:\EBONY'S\US10036729.raw.txt**
Output Set: **N:\CRF4\08052003\J036729.raw**

```
43 CACGGCCGTA GCCCAGTCCG CGACCCCCTC TGTTTCTTCA TCTATTAGCA GCCTCCGGGC   420
44 CGCGACTTCG GGGGCGACTG CCGCCGCCTC CGCCGCCGCA GCCGTCGATA CCGGGTCAGG   480
45 TGGCGGGGGA CAACCCCACG ACACCGCCCC ACGCGGGGCA CGTAAGAAAC AGTAGCCC     538
47 (2) INFORMATION FOR SEQ ID NO: 2:
48       (i) SEQUENCE CHARACTERISTICS:
49           (A) LENGTH: 176 amino acids
50           (B) TYPE: amino acid
51           (C) STRANDEDNESS: single
52           (D) TOPOLOGY: linear
53     (ii) MOLECULE TYPE: peptide
54     (vi) ORIGINAL SOURCE:
55           (A) ORGANISM: Epstein-Barr virus
56     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
57 Met Ala Arg Arg Leu Pro Lys Pro Thr Leu Gln Gly Arg Leu Glu Ala
58 1               5                   10                  15
59 Asp Phe Pro Asp Ser Pro Leu Leu Pro Lys Phe Gln Glu Leu Asn Gln
60             20              25                  30
61 Asn Asn Leu Pro Asn Asp Val Phe Arg Glu Ala Gln Arg Ser Tyr Leu
62         35                  40                  45
63 Val Phe Leu Thr Ser Gln Phe Cys Tyr Glu Glu Tyr Val Gln Arg Thr
64     50                  55                  60
65 Phe Gly Val Pro Arg Arg Gln Arg Ala Ile Asp Lys Arg Gln Arg Ala
66 65                  70                  75                  80
67 Ser Val Ala Gly Ala Gly Ala His Ala His Leu Gly Gly Ser Ser Ala
68                 85                  90                  95
69 Thr Pro Val Gln Gln Ala Gln Ala Ala Ala Ser Ala Gly Thr Gly Ala
70             100                 105                 110
71 Leu Ala Ser Ser Ala Pro Ser Thr Ala Val Ala Gln Ser Ala Thr Pro
72         115                 120                 125
73 Ser Val Ser Ser Ser Ile Ser Ser Leu Arg Ala Ala Thr Ser Gly Ala
74     130                 135                 140
75 Thr Ala Ala Ala Ser Ala Ala Ala Ala Val Asp Thr Gly Ser Gly Gly
76 145                 150                 155                 160
77 Gly Gly Gln Pro His Asp Thr Ala Pro Arg Gly Ala Arg Lys Lys Gln
78             165                 170                 175
80 (2) INFORMATION FOR SEQ ID NO: 3:
81       (i) SEQUENCE CHARACTERISTICS:
82           (A) LENGTH: 1038 base pairs
83           (B) TYPE: nucleic acid
84           (C) STRANDEDNESS: double
85           (D) TOPOLOGY: unknown
86     (ii) MOLECULE TYPE: DNA (genomic)
87     (vi) ORIGINAL SOURCE:
88           (A) ORGANISM: Epstein-Barr virus
89     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
90 ATGCTATCAG GTAACGCAGG AGAAGGAGCA ACAGCCTGCG GAGGTTCGGC CGCCGCGGGC   60
91 CAGGACCTCA TCAGCGTCCC CCGCAACACC TTTATGACAC TGCTTCAGAC CAACCTGGAC   120
92 AACAAACCGC CGAGGCAGAC CCCGCTACCC TACGCGGCCC CGCTGCCCCC CTTTTCCCAC   180
93 CAGGCAATAG CCACCGCGCC TTCCTACGGT CCTGGGGCCG GAGCGGTCGC CCCGGCCGGC   240
```

```
 94 GGCTACTTTA CCTCCCCAGG AGGTTACTAC GCCGGGCCCG CGGGCGGGGA CCCGGGTGCC      300
 95 TTCTTGGCGA TGGACGCTCA CACCTACCAC CCCCACCCAC ACCCCCCTCC GGCCTACTTT      360
 96 GGCTTGCCGG GCCTCTTTGG CCCCCCTCCA CCCGTGCCTC CTTACTACGG ATCCCACTTG      420
 97 CGGGCAGACT ACGTCCCCGC TCCCTCGCGA TCCAACAAGC GGAAAAGAGA CCCCGAGGAG      480
 98 GATGAAGAAG GCGGGGGGCT ATTCCCGGGG GAGGACGCCA CCCTCTACCG CAAGGACATA      540
 99 GCGGGCCTCT CCAAGAGTGT GAATGAGTTA CAGCACACGC TACAGGCCCT GCGCCGGGAG      600
100 ACGCTGTCCT ACGGCCACAC CGGAGTCGGA TACTGCCCCC AGCAGGGCCC CTGCTACACC      660
101 CACTCGGGGC CTTACGGATT TCAGCCTCAT CAAAGCTACG AAGTGCCCAG ATACGTCCCT      720
102 CATCCGCCCC CACCACCAAC TTCTCACCAG GCAGCTCAGG CGCAGCCTCC ACCCCCGGGC      780
103 ACACAGGCCC CCGAAGCCCA CTGTGTGGCC GAGTCCACGA TCCCTGAGGC GGGAGCAGCC      840
104 GGGAACTCTG GACCCCGGGA GGACACCAAC CCTCAGCAGC CCACCACCGA GGGCCACCAC      900
105 CGCGGAAAGA AACTGGTGCA GGCCTCTGCG TCCGGAGTGG CTCAGTCTAA GGAGCCCACC      960
106 ACCCCCAAGG CCAAGTCTGT GTCAGCCCAC CTCAAGTCCA TCTTTTGCGA GGAATTGCTG     1020
107 AATAAACGCG TGGCTTGA                                                    1038
109 (2) INFORMATION FOR SEQ ID NO: 4:
110      (i) SEQUENCE CHARACTERISTICS:
111          (A) LENGTH: 345 amino acids
112          (B) TYPE: amino acid
113          (C) STRANDEDNESS: single
114          (D) TOPOLOGY: linear
115     (ii) MOLECULE TYPE: peptide
116     (vi) ORIGINAL SOURCE:
117          (A) ORGANISM: Epstein-Barr virus
118      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
119 Met Leu Ser Gly Asn Ala Gly Glu Gly Ala Thr Ala Cys Gly Gly Ser
120 1               5                   10                  15
121 Ala Ala Ala Gly Gln Asp Leu Ile Ser Val Pro Arg Asn Thr Phe Met
122             20                  25                  30
123 Thr Leu Leu Gln Thr Asn Leu Asp Asn Lys Pro Pro Arg Gln Thr Pro
124         35                  40                  45
125 Leu Pro Tyr Ala Ala Pro Leu Pro Pro Phe Ser His Gln Ala Ile Ala
126     50                  55                  60
127 Thr Ala Pro Ser Tyr Gly Pro Gly Ala Gly Ala Val Ala Pro Ala Gly
128 65                  70                  75                  80
129 Gly Tyr Phe Thr Ser Pro Gly Gly Tyr Tyr Ala Gly Pro Ala Gly Gly
130             85                  90                  95
131 Asp Pro Gly Ala Phe Leu Ala Met Asp Ala His Thr Tyr His Pro His
132             100                 105                 110
133 Pro His Pro Pro Pro Ala Tyr Phe Gly Leu Pro Gly Leu Phe Gly Pro
134             115                 120                 125
135 Pro Pro Pro Val Pro Pro Tyr Tyr Gly Ser His Leu Arg Ala Asp Tyr
136     130                 135                 140
137 Val Pro Ala Pro Ser Arg Ser Asn Lys Arg Lys Arg Asp Pro Glu Glu
138 145                 150                 155                 160
139 Asp Glu Glu Gly Gly Gly Leu Phe Pro Gly Glu Asp Ala Thr Leu Tyr
140                 165                 170                 175
141 Arg Lys Asp Ile Ala Gly Leu Ser Lys Ser Val Asn Glu Leu Gln His
142             180                 185                 190
143 Thr Leu Gln Ala Leu Arg Arg Glu Thr Leu Ser Tyr Gly His Thr Gly
```

**RAW SEQUENCE LISTING**               DATE: 08/05/2003
PATENT APPLICATION:  **US/10/036,729**      TIME: 10:47:04

Input Set : **N:\EBONY'S\US10036729.raw.txt**
Output Set: **N:\CRF4\08052003\J036729.raw**

```
144          195              200              205
145 Val Gly Tyr Cys Pro Gln Gln Gly Pro Cys Tyr Thr His Ser Gly Pro
146     210              215              220
147 Tyr Gly Phe Gln Pro His Gln Ser Tyr Glu Val Pro Arg Tyr Val Pro
148 225              230              235              240
149 His Pro Pro Pro Pro Pro Thr Ser His Gln Ala Ala Gln Ala Gln Pro
150              245              250              255
151 Pro Pro Pro Gly Thr Gln Ala Pro Glu Ala His Cys Val Ala Glu Ser
152     260              265              270
153 Thr Ile Pro Glu Ala Gly Ala Ala Gly Asn Ser Gly Pro Arg Glu Asp
154     275              280              285
155 Thr Asn Pro Gln Gln Pro Thr Thr Glu Gly His His Arg Gly Lys Lys
156     290              295              300
157 Leu Val Gln Ala Ser Ala Ser Gly Val Ala Gln Ser Lys Glu Pro Thr
158 305              310              315              320
159 Thr Pro Lys Ala Lys Ser Val Ser Ala His Leu Lys Ser Ile Phe Cys
160              325              330              335
161 Glu Glu Leu Leu Asn Lys Arg Val Ala
162              340              345
164 (2) INFORMATION FOR SEQ ID NO: 5:
165     (i) SEQUENCE CHARACTERISTICS:
166         (A) LENGTH: 24 amino acids
167         (B) TYPE: amino acid
168         (C) STRANDEDNESS: single
169         (D) TOPOLOGY: linear
170     (ii) MOLECULE TYPE: peptide
171     (vi) ORIGINAL SOURCE:
172         (A) ORGANISM: Epstein-Barr virus
173     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
174 Ala Val Asp Thr Gly Ser Gly Gly Gly Gly Gln Pro His Asp Thr Ala
175              5              10              15
176 Pro Arg Gly Ala Arg Lys Lys Gln
177              20
179 (2) INFORMATION FOR SEQ ID NO: 6:
180     (i) SEQUENCE CHARACTERISTICS:
181         (A) LENGTH: 30 amino acids
182         (B) TYPE: amino acid
183         (C) STRANDEDNESS: single
184         (D) TOPOLOGY: linear
185     (ii) MOLECULE TYPE: peptide
186     (vi) ORIGINAL SOURCE:
187         (A) ORGANISM: Epstein-Barr virus
188     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
189 Ser Thr Ala Val Ala Gln Ser Ala Thr Pro Ser Val Ser Ser Ser Ile
190              5              10              15
191 Ser Ser Leu Arg Ala Ala Thr Ser Gly Ala Thr Ala Ala Ala
192     20              25              30
194 (2) INFORMATION FOR SEQ ID NO: 7:
195     (i) SEQUENCE CHARACTERISTICS:
```

**RAW SEQUENCE LISTING**           DATE: 08/05/2003
PATENT APPLICATION:   **US/10/036,729**     TIME: 10:47:04

Input Set : **N:\EBONY'S\US10036729.raw.txt**
Output Set: **N:\CRF4\08052003\J036729.raw**

```
196              (A) LENGTH: 15 amino acids
197              (B) TYPE: amino acid
198              (C) STRANDEDNESS: single
199              (D) TOPOLOGY: linear
200        (ii) MOLECULE TYPE: peptide
201        (vi) ORIGINAL SOURCE:
202              (A) ORGANISM: Epstein-Barr virus
203        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
204 Gly Val Pro Arg Arg Gln Arg Ala Ile Asp Lys Arg Gln Arg Ala
205                   5                   10                  15
207 (2) INFORMATION FOR SEQ ID NO: 8:
208        (i) SEQUENCE CHARACTERISTICS:
209              (A) LENGTH: 15 amino acids
210              (B) TYPE: amino acid
211              (C) STRANDEDNESS: single
212              (D) TOPOLOGY: linear
213        (ii) MOLECULE TYPE: peptide
214        (vi) ORIGINAL SOURCE:
215              (A) ORGANISM: Epstein-Barr virus
216        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
217 Gly Gln Pro His Asp Thr Ala Pro Arg Gly Ala Arg Lys Lys Gln
218                   5                   10                  15
220 (2) INFORMATION FOR SEQ ID NO: 9:
221        (i) SEQUENCE CHARACTERISTICS:
222              (A) LENGTH: 12 amino acids
223              (B) TYPE: amino acid
224              (C) STRANDEDNESS: single
225              (D) TOPOLOGY: linear
226        (ii) MOLECULE TYPE: peptide
227        (vi) ORIGINAL SOURCE:
228              (A) ORGANISM: Epstein-Barr virus
229        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
230 Thr Ala Val Ala Gln Ser Ala Thr Pro Ser Val Ser
231                   5                   10
233 (2) INFORMATION FOR SEQ ID NO: 10:
234        (i) SEQUENCE CHARACTERISTICS:
235              (A) LENGTH: 12 amino acids
236              (B) TYPE: amino acid
237              (C) STRANDEDNESS: single
238              (D) TOPOLOGY: linear
239        (ii) MOLECULE TYPE: peptide
240        (vi) ORIGINAL SOURCE:
241              (A) ORGANISM: Epstein-Barr virus
242        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
243 Pro Ser Val Ser Ser Ser Ile Ser Ser Leu Arg Ala
244                   5                   10
246 (2) INFORMATION FOR SEQ ID NO: 11:
247        (i) SEQUENCE CHARACTERISTICS:
248              (A) LENGTH: 12 amino acids
```

**VERIFICATION SUMMARY**          DATE: 08/05/2003
PATENT APPLICATION:  **US/10/036,729**     TIME: 10:47:05

Input Set : **N:\EBONY'S\US10036729.raw.txt**
Output Set: **N:\CRF4\08052003\J036729.raw**

L:19 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:20 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]

**STATISTICS SUMMARY**  DATE: 08/05/2003
PATENT APPLICATION:  **US/10/036,729**  TIME: 10:47:05

Input Set : **N:\EBONY'S\US10036729.raw.txt**
Output Set: **N:\CRF4\08052003\J036729.raw**

Application Serial Number: US/10/036,729
Alpha or Numeric or Xml: Alpha
Application Class:
Application File Date: 12-21-2001
Art Unit: OIPE
Software Application: PatentIN1.0
Total Number of Sequences: 22
Total Nucleotides: 1576
Total Amino Acids: 773
Number of Errors: 0
Number of Warnings: 0
Number of Corrections: 2

**MESSAGE SUMMARY**

**220 C: 2** (Keyword misspelled or invalid format)